STATISTICAL RESEARCH REPORT No 3
Institute of Mathematics June 1974
University of Oslo

COMPARISON OF EXPERIMENTS BY FACTORIZATION

by

Erik N. Torgersen

C O N T E N T

F.0 Summary

Consider random variables $X, Y, \ldots$ whose distributions are known except for an unknown parameter θ belonging to a known finite set Θ. Identify each variable with the experiment it defines and write $X \sim Y$ if X and Y are equally informative. We give first, for given X and Y, a functional criterion for the existence of a Z, indpendent of X, such that $Y \sim (X,Z)$. Combining this with a result on consistent families of experiments, we prove that X has the property that any more informative Y is $\sim (X,Z)$ for some Z independent of X if and only if there is a $\tilde{X} \sim X$ such that:

(i) $\tilde{X}$ is, with probability 1, a non empty sub set of Θ.

(ii) Each θ belongs to some possible value of $\tilde{X}$.

(iii) If $U_1 \neq U_2$ are possible values of $\tilde{X}$ then $\#(U_1 \cap U_2) \leqq 1$.

(iv) If $U_{n+1} = U_1, U_2, \ldots, U_n$ are n possible values of $\tilde{X}$ such that $U_i \cap U_{i+1} \neq \emptyset$; $i = 1, \ldots, n$ then $\bigcap_i U_i \neq \emptyset$.

F.1 Introduction

Consider random variables $X, Y, \ldots$ whose distributions are known except for an unknown parameter θ belonging to a known finite set Θ. One way to increase the information (on θ) given by X is to observe a Z which is independent of X and then consider the combination $Y = (X,Z)$. What about the converse? Suppose Y is more informative than X. When do there exist a Z, independent of X, such that Y is equally informative as (X,Z)? Using essentially the same approach as used by Strassen [4], we find that this is the case if and only if a certain functional associated with Y is everywhere $\leqq$ a certain functional associated with X.

We use this criterion to obtain a description of the variables X having the property that any Y which is more informative than X is equally informative as some pair (X,Z) where Z is independent of X. In doing so we needed an auxiliary result on consistent families of experiments with varying parameter sets which may be of separate interest.

It turns out that X has this property if and only if there is a $\tilde{X}$ such that X and $\tilde{X}$ are equally informative and:

(i) $\tilde{X}$ is, with probability 1, a non empty sub set of Θ.

(ii) Each θ belongs to some possible value of Θ.

(iii) If $U_1 \neq U_2$ are possible values of $\tilde{X}$ then $\#(U_1 \cap U_2) \leqq 1$.

(iv) If $U_1, U_2, \ldots, U_n$ are distinct possible values of $\tilde{X}$ such that $U_i \cap U_{i+1} \neq \emptyset$; $i = 1,2,\ldots,n$ where $U_{n+1} = U_1$ then $\bigcap_{i=1}^{n} U_i \neq \emptyset$.

F.2 Notations, definitions and basic facts.

An experiment will here be defined as a pair $\mathcal{E} = ((\chi, \mathcal{A}), (P_\theta : \theta \in \Theta))$ where $(\chi, \mathcal{A})$ is a measurable space and $(P_\theta : \theta \in \Theta)$ is a family of probability measures on $\mathcal{A}$. $(\chi, \mathcal{A})$ is the sample space of $\mathcal{E}$ while Θ is the parameter set of $\mathcal{E}$.

A notion of "being more informative" for experiments was introduced by Bohnenblust, Shapley and Sherman and may be found in Blackwell [1]. This was generalized by LeCam in [3] (see also Heyer [2]) to the notion of ϵ-deficiency.

Definition. Let $\mathcal{E}$ and $\mathcal{F}$ be experiments with the same parameter set Θ. Then we shall say that $\mathcal{E}$ is more informative than $\mathcal{F}$, if to each decision space $(D, \mathcal{T})$ (i.e. a measurable space) where $\mathcal{T}$ is finite, every bounded loss function $(\theta, d) \hookrightarrow W_\theta(d)$ on $\Theta \times D$ (W_θ is assumed measurable for each $\theta \in \Theta$) and every risk function r obtainable in $\mathcal{F}$ there is a risk function r' obtainable in $\mathcal{E}$ so that

$$r'(\theta) \leqq r(\theta) \quad ; \quad \theta \in \Theta$$

If $\mathcal{E}$ is more informative than $\mathcal{F}$ then we will write this $\mathcal{E} \geqq \mathcal{F}$.

Let $\mathcal{E} = ((\chi, \mathcal{A}), (P_\theta : \theta \in \Theta))$ and $\mathcal{F} = ((\mathcal{Y}, \mathcal{B}), (Q_\theta : \theta \in \Theta))$ be two experiments such that $(P_\theta : \theta \in \Theta)$ is dominated, $\mathcal{Y}$ is a Borel sub set of a complete separable metric space and $\mathcal{B}$ is the class of Borel sub sets of $\mathcal{Y}$. Then it follows from theorem 3 in LeCam's paper [3] (see section 1 in [5]) that $\mathcal{E} \geqq \mathcal{F}$ if and only if there is a Markov kernel M from $(\chi, \mathcal{A})$ to $(\mathcal{Y}, \mathcal{B})$ so that

$$P_\theta M = Q_\theta \quad ; \quad \theta \in \Theta$$

If $\mathcal{E} \geqq \mathcal{F}$ and $\mathcal{F} \geqq \mathcal{E}$ then we shall say that $\mathcal{E}$ and $\mathcal{F}$ are equivalent (or equally informative) and we write this $\mathcal{E} \sim \mathcal{F}$.

All experiments considered in this paper will be assumed, unless otherwise stated, to have the same finite parameter set Θ. We refer to the paper [5] for some of the basic facts on experiments with finite parameter sets.

An experiment $\mathcal{E} = ((\chi, \mathcal{A}), (P_\theta : \theta \in \Theta))$ is equivalent to its standard experiment $\hat{\mathcal{E}} = ((K, \text{Borel class}), (S_\theta : \theta \in \Theta))$ where K is the fundamental probability simplex in R^Θ i.e. $K = \{x : x \in R^\Theta , x \geqq 0$ and $\sum_\theta x_\theta = 1\}$ and S_θ, for each θ, is the probability measure on K induced from P_θ by the map:

$$\left(\frac{dP_\theta}{d\sum_\theta P_\theta} : \theta \in \Theta\right)$$

from $(\chi, \mathcal{A})$ to $(K, \text{Borel class})$. The measure $\sum_\theta S_\theta$ is called the standard measure of $\mathcal{E}$. The standard measure determines the standard experiment since $x \rightsquigarrow x_\theta$ is, for each θ, a version of

$$\frac{dS_\theta}{d\sum_\theta S_\theta} .$$

Two experiments are equivalent if and only if they have the same standard experiments. A measure S on K is the standard measure of same experiment if and only if $\int x_\theta S(dx) = 1 \; ; \; \theta \in \Theta$. If $\mathcal{E}$ and $\mathcal{F}$ are experiments with, respectively, standard measures S and T and $\mathcal{E} \geqq \mathcal{F}$ then we will occasionally write this $S \geqq T$.

Comparison of standard mesures may be expressed in terms of dilatations from K to K i.e. Markov kernels D from K to K so that

$$\int y D(dy|x) = x \; ; \; x \in K .$$

Then it follows (see section 3) directly from the result of LeCam quoted above that an experiment with standard measure S is more informative than an experiment with standard measure T if and only if there is a dilatation D from K to K so that

$$S = DT .$$

Another necessary and sufficient condition for $S \geqq T$ is

$$\int \varphi dS \leqq \int \varphi dT$$

for any continuous concave function φ on K. A direct proof that this implies the existence of a dilatation D such that $S = DT$ is given by Strassen [4]. This may be expressed by inequalities on functionals of $C(K)$ (= the set of continuous functions on K). Let $f \in C(K)$. Then the concave envelope $\hat{f}$ of f is the smallest concave function φ which is $\geqq f$. It follows from the last criterion that

$$S \geqq T \text{ if and only if } \int f dS \leqq \int \hat{f} dT \; ; \; f \in C(K) .$$

The product $\prod_{i=1}^{n} \mathscr{E}_i (= \mathscr{E}_1 \times \mathscr{E}_2 \times \ldots \times \mathscr{E}_n)$ of a finite family $\mathscr{E}_i = ((\chi_i, \mathscr{A}_i)(P_{\theta,i}, \theta \in \Theta))$ of experiments is the experiment $(\Pi(\chi_i, \mathscr{A}_i), (\Pi_i P_{\theta,i} : \theta \in \Theta))$. Products respects equivalence i.e. $\prod_{i=1}^{n} \mathscr{E}_i \sim \prod_{i=1}^{n} \mathscr{F}_i$ when $\mathscr{E}_i \sim \mathscr{F}_i$, $i = 1,2,\ldots n$. Furthermore products are, up to equivalence, commutative and associative. We will occasionally write $S = \prod_{i=1}^{n} S_i = S_1 S_2 \ldots S_n$ when $S, S_1, \ldots, S_n$ are standard measures of, respectively, $\mathscr{E}, \mathscr{E}_1, \ldots, \mathscr{E}_n$ and $\mathscr{E} = \Pi \mathscr{E}_i$. An experiment $\mathscr{E}$ will be called a divisor of the experiment $\mathscr{F}$ if there is an experiment $\mathscr{H}$ (not necessarily unique) so that $\mathscr{E} \times \mathscr{H} \sim \mathscr{F}$.

We will write this $\mathcal{E}|\mathcal{F}$. This notion respects equivalence in the sense that $\tilde{\mathcal{E}}|\tilde{\mathcal{F}}$ when $\mathcal{E}|\mathcal{F}$ and $\mathcal{E}\sim\tilde{\mathcal{E}}$ and $\mathcal{F}\sim\tilde{\mathcal{F}}$. If $\mathcal{E}$ and $\mathcal{F}$ have, respectively, standard measures S and T and $\mathcal{E}|\mathcal{F}$ then we will occasionally write this $S|T$.

Clearly $\mathcal{E}\leqq\mathcal{F}$ when $\mathcal{E}|\mathcal{F}$. Our problem is to determine the experiments $\mathcal{E}$ which permit the converse implication.

A very useful tool for studying products of experiments is the Hellinger transform. The Hellinger transform of an experiment with standard measure S is the map

$$H(\circ|\mathcal{E}) : \quad t \longmapsto \int \prod_\theta x_\theta^{t_\theta} S(dx)$$

from K to $[0,1]$.

This transform determines S, i.e. $\mathcal{E}$ up to equivalence. If $\mathcal{E}_1,\ldots,\mathcal{E}_n$ are experiments then

$$H(\circ|\prod_{i=1}^n \mathcal{E}_i) = \prod_{i=1}^n H(\circ|\mathcal{E}_i) .$$

Clearly $H(t|\mathcal{E}) = 1$ when $t \in \text{ext } K$. We will therefore in the following, when $\# \Theta \geqq 2$, restrict the argument t to the set $\tilde{K} = K - \text{ext } K$ where $\text{ext } K$ is the set of extreme points of K. The extreme points of K are precisely the points $e^\theta : \theta \in \Theta$ where, for each θ,

$$e^\theta_{\theta'} = 1 \text{ or } 0 \text{ as } \theta' = \theta \text{ or } \theta' \neq \theta.$$

If $f \in C(K)$ then $\hat{f}(e^\theta) \underset{\theta}{\equiv} f(e^\theta)$.

The following notations will be used:

$A \subset B$: A is a sub set of B and for some $b \in B$, $b \notin A$.

xy : If $x, y \in R^{\Theta}$ then $xy \in R^{\Theta}$ is defined by
$xy = \{x_\theta y_\theta : \theta \in \Theta\}$

Σx : If $x \in R^{\Theta}$ then $\Sigma x = \sum_\theta x_\theta$

m : denotes the number of elements in Θ

$\# A$: the number of elements in the set A.
Thus $m = \# \Theta$

Θ_x : If $x \in R^{\Theta}$ then $\Theta_x = \{\theta : x_\theta \neq 0\}$

$x \gg y$: If $x, y \in R^{\Theta}$ then we may write $x \gg y$ if $\Theta_x \supseteqq \Theta_y$

e : $e = \sum_\theta e^\theta$ i.e. e is the point in R^{Θ} whose coordinates are all 1.

Measure on K : is synonymuous with "Measure on the class of Borel sub sets of K".

Measure on B : More generally: If B is a Borel sub set of some space then a "measure on B" is used as synonymous with "a measure on the σ-algebra Borel of subsets of B".

supp S : If S is a non negative measure on R^{Θ} then supp S is the smallest closed set F such that $S(\complement F) = 0$.

f_x : If $f \in C(K)$ and $x \in K$ then $f_x \in C(K)$ is defined by:
$f_x(y) = \Sigma xy f(\frac{xy}{\Sigma xy})$; $y \in K$.
If f is concave then f_x is concave.
Clearly $f_x(y) = f_y(x)$ and $f_{\frac{e}{m}} = \frac{1}{m} f$

$\hat{f}_x$: If $f \in C(K)$ and $x \in K$, then $g_x = \hat{h}$ where $g = \hat{f}$ and $h = f_x$. We may thus write $\hat{f}_x$ without worrying about the order of the two operations.

f_S : If $f \in C(K)$ and S is a standard measure on K then $f_S \in C(K)$ is defined by:

$$f_S(x) = \int \Sigma xy f\left(\frac{xy}{\Sigma xy}\right) S(dy) = \int f_x(y) S(dy) \; ; \; x \in K \; .$$

$\hat{f}_S$: $\hat{f}_S = \hat{g}$ where $g = f_S$. This function is, in general, different from h_S where $h = \hat{f}$.

S_x : If S is a standard measure on K and $x \in K$ then S_x is the probability distribution on K whose density w.r.t. S is $y \rightsquigarrow \Sigma xy$.

S_θ : $S_\theta = S_{e_\theta}$; Hence $\{S_\theta : \theta \in \Theta\}$ is the standard experiment whose standard measure is S.

$\mu(f)$: If μ is a measure and f is a function then $\mu(f)$ denotes $\int f d\mu$.

$\mathcal{M}_a$: Any totally informative experiment, i.e. any experiment $((\chi,\mathcal{A}),(P_\theta : \theta \in \Theta))$ where $P_{\theta_1} \wedge P_{\theta_2} = 0$ when $\theta_1 \neq \theta_2$

$\mathcal{M}_i$: Any totally uninformative experiment i.e. any experiment $((\chi,\mathcal{A}),(P_\theta : \theta \in \Theta))$ where P_θ does not depend on θ.

$\mathcal{E}_U$: If $\mathcal{E} = ((\chi,\mathcal{A}),(P_\theta : \theta \in \Theta))$ is an experiment then the restriction $((\chi,\mathcal{A}),(P_\theta : \theta \in U))$ of $\mathcal{E}$ to a sub set U of Θ is denoted by $\mathcal{E}_U$.

F.3 Necessary conditions.

The purpose of this paper is to compare the orderings $S \leqq T$ and $S|T$ for standardmeasures S and T. In his paper [4] Strassen proved the equivalence of ordering by dilatations and by integrals of convex functions within a somewhat more general context than we shall need here. Theorem 2 in [4] implies within our set up.

Theorem F.3.1. (Strassen [4].)

Let $S, T, \ldots$ be standardmeasures. Then:

(i) $\sup\{T(f) : T \geqq S\} = S(\hat{f})$

(ii) $T \geqq S \iff T(f) \leqq S(\hat{f})$; $f \in C(K)$

Using essentially the same idea as in [4] we prove the corresponding results for the ordering "being divisible by"

Theorem F.3.2.

(i) $\sup\{T(f) : T \geqq S\} = m\hat{f}_S(e/m)$; $f \in C(K)$

(ii) $S|T$ if and only if $T(f) \leqq m\hat{f}_S(m^{-1}e)$; $f \in C(K)$

Proof: Let $f \in C(K)$ and suppose $S|T$. Then there is a standardmeasure V so that $SV = T$. Hence

$$T(f) = \int \Sigma xyf\left(\frac{xy}{\Sigma xy}\right)S(dy)V(dx) = \int f_S(x)V(dx) = V(f_S)$$

Hence $\operatorname{Sup}\{T(f) : S|T\} = \sup_V V(f_S) = \sup\{V(f_S) : V \geqq m\delta_{m^{-1}e}\}$ = (by theorem 1) $m\hat{f}_S(m^{-1}e)$.

The last equality may also be seen directly, using only the fact that $(m^{-1}e\ ,\ \hat{f}_S(m^{-1}e))$ belongs to the convex hull of all points $(x,f_S(x))$ where $x \in K$. It follows that the sup is actually obtained by a measure V whose support contains at most $m+1$ points.

Suppose finally that $T(f) \leqq m\hat{f}_S(m^{-1}e)$; $f \in C(K)$.
Let $\mathcal{C}$ denote the set of all standard measures T such that $S|T$. Then $\mathcal{C}$ is a convex and weak* compact sub set of the space $\mathcal{M}$ of finite measures on K. If T did not belong to $\mathcal{C}$ then there is, by Hahn Banach's theorem, a continuous linear functional L on $\mathcal{M}$ such that

$$L(T) > \sup\{L(T') : T' \in \mathcal{C}\}$$

It follows from the continuity of L w.r.t. the weak* topology on $\mathcal{M}$, that it is an evaluation i.e. there is a $f \in C(K)$ so that $L(\mu) = \mu(f)$; $\mu \in \mathcal{M}$. Hence

$$T(f) > \sup\{T'(f) : S|T'\}$$

Thus, by (i),

$$T(f) > m\hat{f}_S(m^{-1}e)$$

i.e. a contradiction. Hence $S|T$. $\square$

We proved also:

<u>Proposition F.3.3</u>

Let $f \in C(K)$ and let S be a standard measure. Then there is a standard measure V whose support contains at most $m+1$ points such that

$$(SV)(f) = m\hat{f}_S(m^{-1}e)$$

Theorems 1 and 2 yield:

<u>Corollary F.3.4</u>

Let $\mathscr{E}$ be an experiment with standard measure S . Then the equivalence

$$\mathscr{F} \geqq \mathscr{E} \iff \mathscr{E} \,|\, \mathscr{F}$$

holds if and only if

$$S(\hat{f}) = m\hat{f}_S(m^{-1}e) \; ; \; f \in C(K)$$

<u>Remark.</u> We have, since $\mathscr{E} \,|\, \mathscr{F} \Rightarrow \mathscr{E} \leqq \mathscr{F}$, always $S(\hat{f}) \geqq m\hat{f}_S(m^{-1}e)$; $f \in C(K)$. It follows that the "=" in the corollary may be replaced with "$\leqq$" .

<u>Proof of the corollary.</u> This follows directly from parts (i) of theorems 1 and 2. $\square$

<u>Corollary F.3.5</u>

Let $\mathscr{E}_i$; $i = 1,2,\ldots,n$ be experiments such that for each i: $\mathscr{F} \geqq \mathscr{E}_i \Rightarrow \mathscr{E}_i \,|\, \mathscr{F}$. Then $\prod_{i=1}^{n} \mathscr{E}_i$ has the same property i.e.:

$$\mathscr{F} \geqq \prod_{i=1}^{n} \quad i \Rightarrow \prod_{i=1}^{n} \mathscr{E}_i \,|\, \mathscr{F}$$

<u>Proof</u>: It suffices to consider the case $n = 2$. Let S_i be the standard measure of $\mathscr{E}_i$; $i = 1,2$. Then S_1S_2 is, be definition, the standard measure of $\mathscr{E}_1 \times \mathscr{E}_2$. Let $f \in C(K)$. Then $f_{S_1S_2} = g_{S_2}$ where $g = f_{S_1}$. On the other hand:

$$(S_1S_2)(\hat{f}) = \int \Sigma xy\hat{f}(\tfrac{xy}{\Sigma xy})S_1(dx)S_2(dy) = \int S_1(\hat{f}_y)S_2(dy)$$

$$= \int m\widehat{(f_y)_{S_1}}(m^{-1}e)S_2(dy) = \int m\widehat{(f_{S_1})_y}(m^{-1}e)S_2(dy)$$

$$= \int m[\hat{f}_{S_1}]_y(m^{-1}e)S_2(dy) = S_2(\hat{f}_{S_1}) = m\hat{g}_{S_2}(m^{-1}e) = m\hat{f}_{S_1S_2}(m^{-1}e) .$$

□

By corollary 4, each $f \in C(K)$ provides a necessary condition for $\mathcal{E} | \mathfrak{F}$. We write this condition out explicity in

Proposition F.3.6

Let S be a standard measure and let $f \in C(K)$. Then $S(\hat{f}) = m\hat{f}_S(m^{-1}e)$ if and only if there is a standard measure V so that for each $x \in \text{supp } S$:

(i) $\int \hat{f}(\frac{xy}{\Sigma xy})V_x(dy) = \hat{f}(x)$

and

(ii) $f(\frac{xy}{\Sigma xy}) = \hat{f}(\frac{xy}{\Sigma xy})$ for V_x almost all y .

Remark. Note that (i) and (ii) holds trivially when $x \in \text{ext } K$.

Proof of the proposition.

1° Suppose V satisfies (i) and (ii) for each $x \in \text{supp } S$. Then, by (i), $S(\hat{f}) = \iint \hat{f}(\frac{xy}{\Sigma xy})V_x(dy)S(dx)$. Hence, using (ii): $S(\hat{f}) = \iint f(\frac{xy}{\Sigma xy})V_x(dy)S(dx) = (SV)(f) = \int f_S(y)V(dy)$

$\leqq \int \hat{f}_S(y)V(dy) \leqq m\hat{f}_S(m^{-1}e)$ where the last "$\leqq$" follows from Jensen's inequality. By the remark: $S(\hat{f}) = m\hat{f}_S(m^{-1}e)$.

2^0 Suppose $S(\hat{f}) = m\hat{f}_S(m^{-1}e)$. By proposition 3 there is a standard measure V so that $SV(f) = m\hat{f}_S(m^{-1}e)$. Hence:

$$S(\hat{f}) = SV(f) = \iint f\left(\frac{xy}{\Sigma xy}\right) V_x(dy)S(dx) = \iint f_y(x)V(dy)S(dx)$$

$\leqq \iint \hat{f}_y(x)V(dy)S(dx) = SV(\hat{f}) \leqq S(\hat{f})$. The last "$\leqq$" follows, since $SV \geqq S$, from the concavity of $\hat{f}$. It follows that we have "=" all the way. In particular S is consentrated on the set

$$\{x : \int f_y(x)V(dy) = \int \hat{f}_y(x)V(dy)\} .$$

This set is, by continuity, closed. Hence
$\int f_y(x)V(dy) = \int \hat{f}_y(x)V(dy)$ when $x \in \text{supp } S$. This proves (ii).
The equality $SV(\hat{f}) = S(\hat{f})$ yield:

$$\iint \hat{f}_y(x)V(dy)S(dx) = \int \hat{f}(x)S(dx)$$

By Jensen's inequality:

$$\int \hat{f}_y(x)V(dy) = \int \hat{f}\left(\frac{xy}{\Sigma xy}\right)V_x(dy) \leqq \hat{f}\left(\int \frac{xy}{\Sigma xy} V_x(dy)\right) =$$

$= \hat{f}(\int xyV(dy)) = \hat{f}(x)$. Hence, by continuity, $\int \hat{f}_y(x)V(dy) = \hat{f}(x)$ when $x \in \text{supp } S$. $\square$

We will now investigate the consequences of the equality $S(\hat{f}) = m\hat{f}_S(m^{-1}e)$ for various functions $f \in C(K)$. Let $\xi \in \text{supp } S - \text{ext } K$. Then there is a $f \in C(K)$ so that $f(e^0) \underset{0}{\equiv} 0$, $f(\xi) = 1$ and $\{x : f(x) = \hat{f}(x)\} = \text{ext } K \cup \{\xi\}$. By proposition 6 there is a V so that:

$$\frac{\xi y}{\Sigma \xi y} \in \text{ext } K \cup \{\xi\} \quad \text{for } V_\xi \text{ almost all } y$$

and

$$\int \hat{f}\left(\frac{\xi y}{\Sigma \xi y}\right)V_\xi(dy) = \hat{f}(\xi) .$$

Hence, since $\hat{f}(e^0) \underset{0}{\equiv} 0$:

$1 = f(\xi) = \hat{f}(\xi) = \hat{f}(\xi)V_\xi(\{y : \frac{\xi y}{\Sigma \xi y} = \xi\})$ so that $\frac{\xi y}{\Sigma \xi y} = \xi$ for V_ξ almost all y, i.e. y is constant on Θ_ξ for V almost all y. This implies (and is implied by) $V_{\theta_1} = V_{\theta_2}$ when $\theta_1, \theta_2 \in \Theta_\xi$.

Consider now any other point $\eta \in \operatorname{supp} S - \operatorname{ext} K$. By proposition 6, $\frac{\eta y}{\Sigma \eta y} \in \operatorname{ext} K \cup \{\xi\}$ for V_η almost all y.

Suppose $\Theta_\xi \not\subset \Theta_\eta$. Then we can't have $\frac{\eta y}{\Sigma \eta y} = \xi$ for any y which is constant on Θ_ξ. Hence $\frac{\eta y}{\Sigma \eta y} \in \operatorname{ext} K$ for V_η almost all y. This is equivalent to $V_{\theta_1} \wedge V_{\theta_2} = 0$ when $\theta_1, \theta_2 \in \Theta_\eta$ and $\theta_1 \neq \theta_2$. Hence $\#(\Theta_\xi \cap \Theta_\eta) \leqq 1$. This implies, for any pair (ξ, η) of support points for S;

$\Theta_\xi \subset \Theta_\eta$ or $\#(\Theta_\xi \cap \Theta_\eta) \leqq 1$.

It follows that $\#(\Theta_\xi \cap \Theta_\eta) \leqq 1$ whenever ξ, η are different points in $\operatorname{supp} S$. We have proved:

<u>Proposition F.3.7</u>

Suppose $S(\hat{f}) = m\hat{f}_S(m^{-1}e)$ whenever $f \in C(K)$ is such that $\#\{x : x \notin \operatorname{ext} K, \hat{f}(x) = f(x)\} \leqq 1$. Then

$$\#(\Theta_\xi \cap \Theta_\eta) \leqq 1$$

for any pair ξ, η of different support points for S. It follows in particular that the map

$$\xi \rightsquigarrow \Theta_\xi$$

from $\operatorname{supp} S$ to the class of sub sets of Θ is 1-1.

<u>Remark.</u> S has, by the last statement, finite support.

Let us proceed to more complicated configurations of points in supp S .

<u>Proposition F.3.8</u>

Suppose $S(\hat{f}) = m\hat{f}_S(m^{-1}e)$ whenever $f \in C(K)$. Let $\xi^1,\xi^2,\ldots,\xi^n$ be $n \geqq 2$ points in supp S such that

(i) $\xi^n \notin \{\xi^1,\ldots,\xi^{n-1}\}$

and

(ii) $\Theta_{\xi^i} \cap \Theta_{\xi^{i+1}} \neq \emptyset$; $i = 1,\ldots,n$ where $\xi^{n+1} = \xi^1$.

Then

$$\Theta_{\xi^{n-1}} \cap \Theta_{\xi^n} = \Theta_{\xi^n} \cap \Theta_{\xi^1}$$

<u>Proof</u>: We may, since the case $n = 2$ is trivial, assume $n \geqq 3$. Consider first the case where $\xi^n \in \text{ext } K$, $\xi^n = e^{\theta_o}$ say. Then $\Theta_{\xi^n} = \{\theta_o\}$. Hence:

$$\Theta_{\xi^{n-1}} \cap \Theta_{\xi^n} = \{\theta_o\} = \Theta_{\xi^n} \cap \Theta_{\xi^1}$$

It follows that we may assume $\xi^n \notin \text{ext } K$.

Let $f \in C(K)$ be such that *)
$\{x : f(x) = \hat{f}(x)\} = \text{ext } K \cup \{\xi^1,\ldots,\xi^{n-1}\}$, $f(x) > 0$

---

*) Let A be any finite sub set of K-ext K . Then there is a $f \in C(K)$ so that $f(e^\theta) \underset{\theta}{\equiv} 0$, $f(a) > 0$ when $a \in A$ and $\{x : f(x) = \hat{f}(x)\} = \text{ext } K \cup A$. This follow from the fact that there are positive numbers $\gamma(a)$; $a \in A$ so that the points (e^θ,o); $\theta \in \Theta$ and $(a,\gamma(a))$; $a \in A$ are the extreme points of the convex hull they span in $K \times R$. This, in turn, may be proved by induction on #A .

when $x \in \{\xi^1,\ldots,\xi^{n-1}\}$-ext K and $f(e^\theta) \underset{\theta}{\equiv} 0$.

By proposition 6 there is a standard measure V so that, for each $x \in \operatorname{supp} S$

$$(\alpha) \quad \int \hat{f}\left(\frac{xy}{\Sigma xy}\right)V_x(dy) = \hat{f}(x)$$

$$(\beta) \quad \frac{xy}{\Sigma xy} \in \text{ext } K \cup \{\xi^1,\ldots,\xi^{n-1}\} \quad \text{for } V_x \text{ almost all } y .$$

Consider a $\xi \in \{\xi^1,\ldots,\xi^{n-1}\}$-ext K .

By (β)

$$\frac{\xi y}{\Sigma \xi y} \in \text{ext } K \cup \{\xi^1,\ldots,\xi^{n-1}\} \quad \text{for } V_\xi \text{ almost all } y .$$

Suppose $\frac{\xi y}{\Sigma \xi y} = \xi^j$ where $1 \leqq j \leqq n-1$. Then

$\Theta_{\xi^j} \subseteqq \Theta_\xi$. By proposition 7

$\xi^j = \xi$ or $\xi^j \in$ ext K . Hence

$\frac{\xi y}{\Sigma \xi y} \in \{\xi\} \cup$ ext K for V_ξ almost all y .

We proceed now as in the proof of the previous proposition

By (α)

$$\hat{f}(\xi) = \int_{\{y:\frac{\xi y}{\Sigma \xi y} = \xi\}} \hat{f}\left(\frac{\xi y}{\Sigma \xi y}\right)V_\xi(dy) = \hat{f}(\xi)V_\xi\left(\left\{y : \frac{\xi y}{\Sigma \xi y} = \xi\right\}\right)$$

so that

$$V_\xi\left(\left\{y : \frac{\xi y}{\Sigma \xi y} = \xi\right\}\right) = 1$$

Hence y is constant on Θ_ξ for V almost all y . It follows that $V_{\theta_1} = V_{\theta_2}$ when $\theta_1, \theta_2 \in \Theta_\xi$. This conclusion is trivial if

$\xi \in \text{ext } K$. Hence, by (ii), $V_{0_1} = V_{0_2}$ when $\theta_1, 0_2 \in \bigcup_{i=1}^{n-1} \Theta_{\xi^i}$.

Consider next ξ^n . By (ii)

$\frac{\xi^n y}{\Sigma \xi^n y} \in \text{ext } K \cup \{\xi^1, \ldots, \xi^{n-1}\}$ for V_{ξ^n} almost all y .

Hence $V_{0_1} \wedge V_{0_2} = 0$ when $0_1, 0_2 \in \Theta_{\xi^n}$ and $0_1 \neq 0_2$. There are, by proposition 7, points 0_n and 0_{n-1} in Θ so that $\Theta_{\xi^{n-1}} \cap \Theta_{\xi^n} = \{0_{n-1}\}$ and $\Theta_{\xi^n} \cap \Theta_{\xi^1} = \{0_n\}$. Suppose $0_{n-1} \neq 0_n$. Then, since $0_{n-1}, 0_n \in \Theta_{\xi^n}$, $V_{0_{n-1}} \wedge V_{0_n} = 0$. On the other hand, since $0_{n-1}, 0_n \in \bigcup_{i=1}^{n-1} \Theta_{\xi^i}$, $V_{0_{n-1}} = V_{0_n}$ i.e. a contradiction.

Hence $0_{n-1} = 0_n$. □

Corollary F.3.9

Suppose $S(\hat{f}) = m\hat{f}_S(m^{-1}e)$; $f \in C(K)$. Let $\xi^1, \ldots, \xi^n$ be n distinct points in supp S such that $\Theta_{\xi^i} \cap \Theta_{\xi^{i+1}} \neq \emptyset$; $i = 1, 2, \ldots, n$ where $\xi^{n+1} = \xi^1$. Then $\bigcap_{i=1}^{n} \Theta_{\xi_i} \neq \emptyset$.

Proof: The corollary is trivial if $n \leqq 2$. Suppose the statement holds whenever $n < q$ where $q \geqq 3$. Let $\xi^1, \ldots, \xi^q$ be q distinct points in supp S such that $\Theta_{\xi^i} \cap \Theta_{\xi^{i+1}} \neq \emptyset$; $i=1, \ldots, q$ where $\xi^{q+1} = \xi^1$. By proposition 7 there are points $0_1, \ldots, 0_q$ in Θ so that

$$\Theta_{\xi^i} \cap \Theta_{\xi^{i+1}} = \{0_i\} .$$

By proposition 8, $0_{q-1} = 0_q$. Hence $\Theta_{\xi^{q-1}} \wedge \Theta_{\xi^1} = \{0_{q-1}\}$.

Hence, by the induction hypothesis and proposition 6 ,

$\bigcap_{i=1}^{q-1} \Theta_{\xi^i} = \{0_{q-1}\}$ i.e. $0_1=0_2=\ldots=0_{q-1}$ so that $0_1=\ldots=0_q$. It follows that $\bigcap_{i=1}^{q} \Theta_{\xi^i} = \{0_q\}$. □

It follows from these results that we must examine classes $\mathcal{U}$ of sub sets of Θ which satisfies:

A_1 : $\cup\{U : U \in \mathcal{U}\} = \Theta$ and $\emptyset \notin \mathcal{U}$

A_2 : $U_1, U_2 \in \mathcal{U}$ and $U_1 \neq U_2 \Rightarrow \#(U_1 \cap U_2) \leqq 1$

A_3 : If $U_1, U_2, \ldots, U_n$ are n distinct members of $\mathcal{U}$ such that $U_i \cap U_{i+1} \neq \emptyset$; $i = 1,\ldots,n$ where $U_{n+1} = U_1$ then $\bigcap_{i=1}^{n} U_i \neq \emptyset$.

Let $\mathcal{U}$ satisfy A_1, A_2 and A_3 . Occasionally we may have sets $U_1,\ldots,U_n$ in $\mathcal{U}$, not necessarily distinct, such that $U_i \cap U_{i+1} \neq \emptyset$; $i = 1,2,\ldots,n$ where $U_{n+1} = U_1$. Then we are, in general, not permitted to conclude $\bigcap_{i=1}^{n} U_i \neq \emptyset$. We have, however, $\bigcap_{i=1}^{r} V_i \neq \emptyset$ where $V_1,\ldots,V_r$ are obtained from $U_1,\ldots,U_n$ by the following reduction procedure:

Put $V_i = U_i$ as long as $\#\{U_1,\ldots,U_i\} = i$. If $\#\{U_1,\ldots,U_n\} = n$ then $r = n$ and $V_i = U_i$; $i = 1,\ldots,n$ and there is nothing to prove. If $\#\{U_1,\ldots,U_n\} < n$ then there is a largest j so that $\#\{U_1,\ldots,U_j\} = j$. Clearly $1 \leqq j \leqq n-1$. By assumption there is a smallest i in $\{1,2,\dots j\}$ so that $U_i = U_{j+1}$. Consider the reduced sequence:

$$U_1,\dots,U_i \quad U_{j+2},U_{j+3},\dots,U_n \quad \text{if} \quad j \leqq n-2$$

and

$$U_1,\dots,U_i \quad \text{if} \quad j = n-1 .$$

Continuing this procedure we obtain eventually a sequence $V_1,\dots,V_r$ of distinct sets in $\{U_1,\dots,U_n\}$ such that $V_1 = U_1$, $V_r = U_n$, $V_i \cap V_{i+1} \neq \emptyset$ $\quad i = 1,\dots,r$ where $V_{r+1} = V_1$.

By assumption A_3, $\bigcap_{i=1}^{r} V_i \neq \emptyset$.

F.4 An auxiliary result on consistent families of experiments

In order to establish sufficient conditions we shall need the following result on consistent families of experiments.

Theorem F.4.1

Let $\mathcal{U}$ be a class of subsets of Θ satisfying A_1, A_2 and A_3 of section 3 and let $\{\mathcal{H}^U : U \in \mathcal{U}\}$ be a family of experiments such that $\mathcal{H}^U$, for each U, has parameter set U. Then there is an experiment $\mathcal{H}$ with parameter set Θ so that $\mathcal{H}^U$, for each $U \in \mathcal{U}$, is equivalent to the restriction of $\mathcal{H}$ to U.

The proof is based on:

Proposition*) F.4.2

Let $\mathcal{E}$ be any experiment with parameter set Θ and let $\theta_0 \in \Theta$. Then there are atomless measures $P_\theta : \theta \in \Theta$ on $[0,1]$ so that $\mathcal{E}$ is equivalent to $(P_\theta : \theta \in \Theta)$ and P_{θ_0} is the uniform distribution on $[0,1]$.

Proof: We may, since the case $\# \Theta = 1$ is trivial assume $\# \Theta \geqq 2$. There are, since $\mathcal{E}$ is equivalent to its standard experiment and K is Borel isomorph to $[-1,0[$ probability measures $Q_\theta : \theta \in \Theta$ on $[-1,0[$ such that $\mathcal{E}$ is equivalent to $(Q_\theta : \theta \in \Theta)$. Put $A = \{x : \sum_\theta Q_\theta(x) > 0\}$. Then A is an enumerable subset of $[-1,0[$. Let $a_0, a_1, a_2, \ldots$ be an enumeration of its elements. Put $\mathcal{Y} = [-1,0[\cup \cup\{[\nu,\nu+1[: a_\nu$ is defined $\}$ Then $\mathcal{Y} = [-1,\infty[$ or $= [1,n+1[$ as $\# A = \infty$ or $\# A = n+1$. Let λ_ν denote the uniform distribution on $[\nu,\nu+1[$ and put

*) Actually the proposition carries over, with only minor changes in the proof, to the case of infinite parameter sets provided the experiment is separable.

$$V_\theta(B) = Q_\theta(B \cap A^c) + \sum_\nu Q_\theta(a_\nu)\, r_\nu\,(B \cap [\nu, \nu+1])$$

for each Borel subset B of $\mathcal{Y}$.

It is easily checked that $\|\Sigma c_\theta Q_\theta\| = \|\Sigma c_\theta V_\theta\|$ for each $c \in R^\Theta$. By the corollary to proposition 12 in Le Cam's paper [3] the experiment $\mathcal{E}$ is equivalent to $\{\nu_\theta : \theta \in \Theta\}$. [Alternatively we may, Torgersen [5], conclude that the experiments are equivalent for testing problems and, consequently, equivalent.] The measures $V_\theta : \theta \in \Theta$ are obviously continuous; i.e. atomless. Let $V_{\theta,a}$ and $V_{\theta,s}$ denote, respectively, the V_{θ_0} absolutely continuous part of V_θ and the V_{θ_0} singular part of V_θ. Hence

$$V_\theta = V_{\theta,a} + V_{\theta,s}$$

$$V_{\theta,a} \ll V_{\theta_0} \quad \text{and} \quad V_{\theta,s} \wedge V_{\theta_0} = 0$$

$$V_{\theta_0,a} = V_{\theta_0}, \quad V_{\theta_0,s} = 0$$

Put $\tau_\theta = \|V_{\theta,a}\|$; $\theta \in \Theta$. Then $\tau_{\theta_0} = 1$.

Put $F(x) = V_{\theta_0}(]-\infty,x[)$; $x \in R$ and let, for each $p \in]0,1[$, $F^{-1}(p)$ be the lower p fractile of V_{θ_0} i.e. $F^{-1}(p) = \inf\{x : F(x) \geqq p\}$. Put $F_{\theta,a}(x) = V_{\theta,a}(]-\infty,x[)$; $x \in R$ and define, for each θ, the distribution function K_θ on $]0,1[$ by:

$$K_\theta(p) = F_{\theta,a}(F^{-1}(p)) \ ; \quad p \in]0,1[$$

Then K_θ is continuous, $K_\theta(0+) = 0$ and $K_\theta(1-) = \tau_\theta$. It follows that K_θ is the distribution function of a non negative and atomless measure μ_θ on $[0,1]$. The total mass of the measure μ_θ is $K_\theta(1-) = \tau_\theta$. The measures $\mu_\theta : \theta \in \Theta$ have the following properties:

(α) μ_{θ_0} is the uniform distribution on $[0,1]$

(β) $\mu_{\theta_0} \gg \mu_\theta : \theta \in \Theta$

[Here is a proof of (β): Choose a $\theta \in \Theta$ and let $\varepsilon > 0$. Then there is, since $V_{\theta_0} \gg V_{\theta,a} : \theta \in \Theta$ a $\delta > 0$ so that $\sum_{i=1}^{n} |F_{\theta,a}(y_i) - F_{\theta,a}(x_i)| < \varepsilon$ when $x_1 \leqq y_1 \leqq x_2 \leqq \ldots \leqq y_{n-1} \leqq x_n \leqq y_n$ and $\sum_{i=1}^{n} |F(y_i) - F(x_i)| < \delta$

Let $p_1 \leqq q_1 \leqq p_2 \leqq \ldots \leqq q_{n-1} \leqq p_n \leqq q_n$ and $\sum_{i=1}^{n} |q_i - p_i| < \delta$. Put $x_i = F^{-1}(p_i)$ and $y_i = F^{-1}(q_i)$. Then $\sum_{i=1}^{n} |F(y_i) - F(x_i)| = \sum_{i=1}^{n} |q_i - p_i| < \delta$. Hence $\sum_{i=1}^{n} |K_\theta(q_i) - K_\theta(p_i)| = \sum_{i=1}^{n} |F_{\theta,a}(y_i) - F_{\theta,a}(x_i)| < \varepsilon$]

Put $X = F^{-1}$ and $Y = F$. Then $\mu_\theta X^{-1} = V_{\theta,a}$; $\theta \in \Theta$ and $V_{\theta,a} Y^{-1} = \mu_\theta$; $\theta \in \Theta$. Hence

$$\|\Sigma c_\theta \mu_\theta\| = \|\Sigma c_\theta V_{\theta,a}\| \quad \text{when} \quad c \in R^\Theta.$$

Let J be a Borel subset of $]0,1[$ which has Lebesgue measure zero and is Borel isomorph to R by the map $\rho : R \to J$. Put $\sigma_\theta = V_{\theta,s}\rho^{-1}$. Clearly $\|\Sigma c_\theta \sigma_\theta\| = \|\Sigma c_\theta V_{\theta,s}\|$; $c \in R^\Theta$. In particular $\|\sigma_\theta\| = \|V_{\theta,s}\| = 1 - \tau_\theta$. The proof follows now from the same result of Le Cam by putting $P_\theta = \mu_\theta + \sigma_\theta$ and noting that:

$$\|\Sigma c_\theta P_\theta\| = \|\Sigma c_\theta \mu_\theta\| + \|\Sigma c_\theta \sigma_\theta\| = \|\Sigma c_\theta V_{\theta,a}\| + \|\Sigma c_\theta V_{\theta,s}\|$$
$$= \|\Sigma c_\theta V_\theta\| = \|\Sigma c_\theta Q_\theta\|.$$

□

Corollary F.4.3

Let $\mathcal{E}'$ and $\mathcal{E}''$ be experiments with parameter sets, respectively, Θ' and Θ''. Suppose $\Theta' \cup \Theta'' = \Theta$ and $\#(\Theta' \cap \Theta'') \leqq 1$. Then there is an experiment $\mathcal{E}$ with parameter set Θ such that

$$\mathcal{E}' \sim \mathcal{E}_{\Theta'} \quad \text{and} \quad \mathcal{E}'' \sim \mathcal{E}_{\Theta''}.$$

<u>Proof:</u> Let $\mathcal{E}' \sim (P'_\theta : \theta \in \Theta')$ and $\mathcal{E}'' \sim (P''_\theta : \theta \in \Theta'')$. We may by proposition 2, assume that $\mathcal{E}'$ and $\mathcal{E}''$ have the same samplespaces and that $P'_\theta = P''_\theta$ when $\theta \in \Theta' \cap \Theta''$. Put $P_\theta = P'_\theta$ or $= P''_\theta$ as $\theta \in \Theta'$ or $\theta \in \Theta''$. Then the experiment $\mathcal{E} = (P_\theta : \theta \in \Theta)$ has the desired properties.

The corollary is actually a particular case of the theorem. We shall now prove the theorem by generalizing the patching used in the proof of the corollary.
Condition A.3 ensures that the patching is consistent.

<u>Proof of theorem 1</u>

Denote by $\mathcal{C}$ the class of subsets C of Θ which have the property that there is an experiment with parameter set C such that the restriction of this experiment to an $U \in \mathcal{U}$ which is contained in C is equivalent to $\mathcal{H}^U$.
Introduce an equivalence relation § or Θ by θ' § θ'' if and only if there is a sequence $U_1, U_2, \ldots, U_n$ in $\mathcal{U}$ such that $\theta' \in U_1$, $\theta'' \in U_n$ and $U_i \cap U_{i+1} \neq 0$ when $1 \leqq i < n$. Let $\Theta_t : t \in T$ be the equivalence classes and suppose we have proved the theorem when $\# T = 1$. Put $\mathcal{U}_t = \{U : U \in \mathcal{U}$ & $U \subseteqq \Theta_t\}$. Clearly $\mathcal{U}_t$ satisfies A_1, A_2, A_3 w.r.t. Θ_t. Furthermore there is only one equivalence class w.r.t. the equivalence relation induced by $\mathcal{U}_t$ on Θ_t. It follows that there is, for each t, an experiment $\mathcal{F}^{(t)}$ with parameter set Θ_t and sample space $[0,1]$ such that $\mathcal{F}^{(t)}_U \sim \mathcal{H}^U$ where $U \subseteqq \Theta_t$. If $\mathcal{F}^{(t)} = (P_{\theta,t} : \theta \in \Theta_t)$ then we may put $\mathcal{H} = (P_\theta : \theta \in \Theta)$ where $P_\theta = P_{\theta,t}$ when $\theta \in \Theta_t$. We may therefore assume that there is only one equivalence class.

Choose next any $\tilde{U} \in \mathcal{U}$ and define recursively subsets $\Theta_1, \Theta_2, \ldots$ of Θ by

$$\Theta_1 = \tilde{U}$$

$$\Theta_{i+1} = \cup\{U : U \in \mathcal{U}, U \cap \Theta_i \neq \emptyset\}$$

By A.1: $\Theta_1 \subseteqq \Theta_2 \subseteqq \ldots \subseteqq \Theta$ and, since there are only one equivalence class, $\bigcup_i \Theta_i = \Theta$. It follows that $\Theta_i = \Theta$ when i is sufficiently large. It is, by induction, seen that $\theta \in \Theta_i$ if and only if there are sets $U_1 = \tilde{U}, U_2, \ldots, U_i$ in $\mathcal{U}$ so that $U_i \ni \theta$ and $U_1 \cap U_2 \neq \emptyset$, $U_2 \cap U_3 \neq \emptyset, \ldots, U_{i-1} \cap U_i \neq \emptyset$.

Put $I = \{i : \Theta_i \in \mathcal{C}\}$. It follows from A.2 that $1 \in I$ and trivially, $i \in I$ if $1 \leqq i \leqq j \in I$. We will be through if we can prove:

Statement: Let $i \in I$. Then $i+1 \in I$.

Proof of the statement: If $\Theta_i = \Theta_{i+1}$ then there is nothing to prove. Suppose $\Theta_i \subset \Theta_{i+1}$. Then $\Theta_{i+1} = \Theta_i \cup \bigcup_{i=1}^{r} V_i$ where the distinct sets $V_1, V_2, \ldots, V_r$ are the sets in $\mathcal{U}$ which intersects Θ_i without being contained in Θ_i. Put $\Theta^{(o)} = \Theta_i$ and $\Theta^{(j)} = \Theta_i \cup V_1 \cup V_2 \cup \ldots \cup V_j$; $1 \leqq j \leqq r$. Consider the set $J = \{j : 0 \leqq j \leqq r, \ \Theta^{(j)} \in \mathcal{C}\}$. Clearly $0 \in J$ and $j' \in J$ when $1 \leqq j' \leqq j'' \in J$. Suppose $\theta', \theta'' \in V_j \cap \Theta_i$. By the definition of Θ_i there are distinct subsets $U_1^*, U_2^*, \ldots, U_a^*$ of Θ_i so that $U_1^* \cap U_2^* \neq \emptyset$, $U_2^* \cap U_3^* \neq \emptyset, \ldots, U_{a-1}^* \cap U_a^* \neq \emptyset$, $\theta' \in U_1^*$, $\theta'' \in U_a^*$. By A.3

$$U_1^* \cap U_2^* \cap \ldots \cap U_a^* \cap V_j \neq \emptyset.$$

Hence

$$\{\theta'\} = U_1^* \cap V_j = U_1^* \cap U_2^* \cap \ldots \cap U_a^* \cap V_j = U_a^* \cap V_j = \{\theta''\}$$

so that

$$\theta' = \theta''$$

It follows that

$$\#(V_j \cap \Theta_i) = 1 \; ; \quad j = 1,2,\ldots,r$$

Consider next distinct indexes j', j'' in $\{1,2,\ldots,r\}$. Let $\theta \in V_{j'} \cap V_{j''}$. Then, by assumption A_2, $V_{j'} \cap V_{j''} = \{\theta\}$. As we have seen $\#(V_{j'} \wedge \Theta_i) = \#(V_{j''} \wedge \Theta_i) = 1$. It follows, using A.2 once more, that there are distinct elements θ', θ'' in Θ so that

$$V_{j'} \cap \Theta_i = \{\theta'\}$$

and

$$V_{j''} \cap \Theta_i = \{\theta''\}$$

There are, by the definition of Θ_i, distinct subsets $U_{(1)},\ldots,U_{(d)}$ of Θ_i such that $U_{(1)},\ldots,U_{(d)} \in \mathcal{U}$, $U_{(k)} \cap U_{(k+1)} \neq \emptyset$ when $1 \leqq k < d$ and $\theta' \in U_{(1)}$ while $\theta'' \in U_{(d)}$. By A.3 ; $V_{j'} \cap U_{(1)} \cap U_{(2)} \cap \ldots \cap U_{(d)} \cap V_{j''} \neq \emptyset$. Hence $\{\theta\} = V_{j'} \cap U_{(1)} \cap \quad \cap U_{(d)} \cap V_{j''}$ so that $\theta \in \Theta_i$. This proves that

$$V_{j'} \cap V_{j''} \subseteqq \Theta_i \qquad \text{when} \quad j' \neq j''$$

Combining the last two results we find that $\#(\Theta^{(j)} \cap V_{j+1}) \leqq 1$. Hence, by the corollary, $j+1 \in J$ when $j \in J$ and $j < r$. It follows, by induction that $r \in J$ i.e. $\Theta_{i+1} \in \mathcal{C}$. This proves the statement and, consequently, completes our proof of the theorem. $\square$

The theorem yield the following divisibility criterion:

<u>Corollary F.4.4</u>

Let $\mathcal{U}$ be a class of subsets of Θ satifying A_1, A_2 and A_3. Suppose $\mathcal{E} = ((\chi, A),(P_\theta : \theta \in \Theta))$ and $\mathcal{F} = ((\mathcal{Y}, \mathcal{B}),(Q_\theta : \theta \in \Theta))$ are experiments with parameter sets Θ such that:

(i) $\mathcal{E}_U \mid \mathcal{F}_U \; ; \quad U \in \mathcal{U}$

(ii) $\bigwedge_F P_\theta = \bigwedge_F Q_\theta = 0$ when $\emptyset \subset F \subsetneqq U \; ; \quad U \in \mathcal{U}$

Then $\mathcal{E} | \mathcal{F}$

Remark. (ii) is automatically satisfied when $\mathcal{E} \leqq \mathcal{F}$.

Proof of the corollary:

By (i) there is, for each U , an experiment $\mathcal{H}^U$ with parameter set U such that

$$\mathcal{E}_U \times \mathcal{H}^U \sim \mathcal{F}_U$$

By the theorem there is an experiment $\mathcal{H}$ with parameter set Θ so that

$$\mathcal{H}_U \sim \mathcal{H}^U \; ; \quad U \in \mathcal{U} \; .$$

Hence

$$\mathcal{E}_U \times \mathcal{H}_U \sim \mathcal{F}_U \qquad U \in \mathcal{U}$$

so that

$$H(t|\mathcal{E}) \cdot H(t|\mathcal{H}) = H(t|\mathcal{F})$$

when $\Theta_t \subseteqq U \in \mathcal{U}$.

By (ii)

$$H(t|\mathcal{E}) = H(t|\mathcal{F}) = 0 \quad \text{when} \quad \Theta_t \nsubseteqq U \; ; \quad U \in \mathcal{U}$$

Hence

$$H(t|\mathcal{E}) \, H(t|\mathcal{H}) = H(t|\mathcal{F}) \; ; \quad t \in K$$

i.e. $\mathcal{E} \times \mathcal{H} \sim \mathcal{F}$. $\square$

F.5 Characterization of the experiments $\mathcal{E}$ which permits the implication $\mathcal{E} \leqq \mathcal{F} \Rightarrow \mathcal{E} | \mathcal{F}$

Let us begin by considering some simple experiments which all have the desired property.

To each $a \in [0,1]^\Theta$ we associate an experiment $\mathcal{S}_a = ((\chi,\mathcal{A}),(P_\theta : \theta \in \Theta))$ where $\chi = \{U : U \subseteqq \Theta$ and $\# U = 1$ or $\# \Theta\}$, $\mathcal{A}$ = the class of all subsets of χ and P_θ assigns masses a_θ and $1 - a_\theta$ to, respectively, Θ and $\{\theta\}$. The Hellinger transform of $\mathcal{S}_a$ is simply:

$$H(t|\mathcal{S}_a) = \prod_\theta a_\theta^{t_\theta} \quad ; \quad t \in K - \operatorname{ext} K .$$

It follows that:

$$\mathcal{S}_a \times \mathcal{S}_b \sim \mathcal{S}_{ab} \quad ; \quad a,b \in [0,1]^\Theta$$

Restrictions of these experiments to subsets of the parameter set is of the same type. More specifically $[\mathcal{S}_a]_U = \mathcal{S}_{a|U}$ when U is a non empty subset of Θ and $a|U$ is the restriction of a to U. This class includes the experiments $\mathcal{M}i$ and $\mathcal{M}_a$ since:

$$\mathcal{S}_a \sim \mathcal{M}_a \iff \# \Theta_a \leqq 1$$

and

$$\mathcal{S}_a \sim \mathcal{M}i \iff a = e .$$

Within the class $\{\mathcal{S}_a : a \in R^\Theta\}$ the ordering "being more informative" coincides with the ordering "being divisible by". This follows from

Proposition F.5.1

Let $a,b \in [0,1]^\Theta$. Then $a \geqq b \Rightarrow \mathcal{S}_a | \mathcal{S}_b \iff \mathcal{S}_a \leqq \mathcal{S}_b$ where the first "$\Rightarrow$" may be replaced by equivalence "$\iff$" provided $\mathscr{S}_b \not\sim \mathscr{M}a$.

<u>Proof</u>: Suppose $a \geq b$. Then $b = ac$ where $c_\theta = b_\theta / a_\theta$ when $a_\theta > 0$. Hence $\mathscr{S}_b \sim \mathscr{S}_a \times \mathscr{S}_c$ so that $\mathscr{S}_a \mid \mathscr{S}_b$. The implication $\mathscr{S}_a \mid \mathscr{S}_b \Rightarrow \mathscr{S}_a \leq \mathscr{S}_b$ is trivial. Suppose next that $\mathscr{S}_a \leq \mathscr{S}_b \not\sim \mathscr{M}a$. Then $\mathscr{S}_a \not\sim \mathscr{M}a$. The standard measures of $\mathscr{S}_a$ and $\mathscr{S}_b$ assigns, respectively, masses $1-a_\theta$ and $1-b_\theta$ to e^θ. Hence $1-a_\theta \leqq 1-b_\theta$; $\theta \in \Theta$ i.e. $a \geqq b$. □

<u>Corollary F.5.2</u>

Let $a,b \in [0,1]^\Theta$. Then $\mathscr{S}_a \sim \mathscr{S}_b \not\sim \mathscr{M}a \iff a = b$ and $\# \Theta_b \geqq 2$.

<u>Proof</u>: This follows directly from the proposition. □

Factors of the form $\mathscr{S}_a$ may be determined from:

<u>Proposition F.5.3</u>

Let $\mathscr{E} = ((\chi,\mathscr{A}),(P_\theta : \theta \in \Theta))$ be an experiment and let, for each θ, $\tilde{P}_\theta$ be the $\sum_{\theta' \neq \theta} P_{\theta'}$ absolutely continuous part of P_θ. Put $Q_\theta = \|\tilde{P}_\theta\|^{-1}\tilde{P}_\theta$ or $= P_\theta$ as $\tilde{P}_\theta \neq 0$ or $\tilde{P}_\theta = 0$, and let $\mathscr{F}$ denote the experiment $\mathscr{F} = ((\chi,\mathscr{A}),(Q_\theta : \theta \in \Theta))$. Then

$$\mathscr{E} \sim \mathscr{S}_a \times \mathscr{F}$$

if and only if $a_\theta = \|\tilde{P}_\theta\|$ when $\tilde{P}_\theta \neq 0$.

<u>Proof</u>: Put $\mu = \sum_\theta P_\theta$ and $f_\theta = dP_\theta/d\mu$. Then $H(t|\mathscr{E}) = \int \Pi f_\theta^{t_\theta} d\mu$. Furthermore:

$d\tilde{P}_\theta/d\mu = f_\theta I_{f \in \tilde{K}}$ where $\tilde{K} = K - \operatorname{ext} K$. Put $U = \{\theta : \tilde{P}_\theta \neq 0\}$. Then $H(t|\mathscr{F}) = \int \Pi_U [\|\tilde{P}_\theta\|^{-1} f_\theta I_{f \in \tilde{K}}]^{t_\theta} \cdot \Pi_{U^c} f_\theta^{t_\theta} d\mu = H(t|\mathscr{E}) / \Pi_U \|\tilde{P}_\theta\|^{t_\theta}$. Hence

$H(t|\mathscr{E}) = \prod_{\theta \in U} \|\tilde{P}_\theta\|^{t_\theta} H(t|\mathscr{F})$; $t \in \tilde{K}$. Suppose $a_\theta = \|\tilde{P}_\theta\|$ when $\|\tilde{P}_\theta\| > 0$. Then $H(t|\mathscr{E}) = \prod_\theta a_\theta^{t_\theta} H(t|\mathscr{F})$; i.e. $\mathscr{E} \sim \mathscr{I}_a \times \mathscr{F}$.

(The values of a_θ when $\theta \notin U$ does not matter since $H(t|\mathscr{E}) = H(t|\mathscr{F}) = 0$ when $\Theta_t \not\subseteq U$)

Conversely, suppose $\mathscr{E} \sim \mathscr{I}_b \times \mathscr{F}$ i.e. $H(t|\mathscr{E}) = \prod_\theta b_\theta^{t_\theta} H(t|\mathscr{F})$.

Let $\theta \in U$. Then $\tilde{P}_\theta \neq 0$. It follows that there is a $\theta' \neq \theta$ so that $P_{\theta'} \wedge P_\theta \neq 0$. Put $t_\theta = \lambda$, $t_{\theta'} = 1 - \lambda$ where $\lambda \in]0,1[$.

Then $0 < H(t|\mathscr{E}) = b_\theta^\lambda b_{\theta'}^{1-\lambda} H(t|\mathscr{F})$ and

$$0 < H(t|\mathscr{E}) = a_\theta^\lambda a_{\theta'}^{1-\lambda} H(t|\mathscr{F})$$

Hence $0 < b^\lambda b_{\theta'}^{1-\lambda} = a_\theta^\lambda a_{\theta'}^{1-\lambda}$. $\lambda \to 1$ yield

$$b_\theta = a_\theta$$

This proves the "if and only if". $\square$

<u>Corollary F.5.4</u>

Let $\mathscr{E} = ((\chi, \mathcal{A}), (P_\theta : \theta \in \Theta))$ be an experiment and let $a \in [0,1]^\Theta$. Denote by $\tilde{P}_\theta$ the $\sum_{\theta' \neq \theta} P_{\theta'}$ absolutely continuous part of P_θ. Then the following conditions are equivalent:

(i) $a_\theta \geqq \|\tilde{P}_\theta\|$; $\theta \in \Theta$

(ii) $\mathscr{I}_a | \mathscr{E}$

(iii) $\mathscr{I}_a \leqq \mathscr{E}$

<u>Proof</u>: Put $b_\theta = \|\tilde{P}_\theta\|$; $\theta \in \Theta$. By the proppsition $\mathscr{E} \sim \mathscr{I}_b \times \mathscr{F}$ where $\mathscr{F}$ is defined as in the proposition. Suppose $a_\theta \geqq b_\theta$; $\theta \in \Theta$. By proposition 1, $\mathscr{I}_a | \mathscr{I}_b$. Hence, since $\mathscr{I}_b | \mathscr{E}$, $\mathscr{I}_a | \mathscr{E}$. Thus (i) ==> (ii). It remains, since (ii) ==> (iii) is trivial, to prove (iii) ==> (i). Suppose $\mathscr{I}_a \leqq \mathscr{E}$. We may,

since (i) is trivial when $\mathcal{E} \sim \mathcal{M}_a$, assume $\mathcal{E} \not\sim \mathcal{M}_a$. Let S and T denote, respectively, the standard measures of $\mathcal{J}_a$ and $\mathcal{E}$. Then $S(e^\theta) \leqq T(e^\theta)$ i.e. $1-a_\theta \leqq 1-b_\theta$. Hence $a \geqq b$. □

Not all experiments $\mathcal{E}$ permitting the implication "$\mathcal{E} \leqq \mathcal{F} \Rightarrow \mathcal{E} \mid \mathcal{F}$" belongs, up to equivalence, to the class $\{\mathcal{J}_a : a \in [0,1]^\Theta\}$. We must therefore, in order to obtain the complete chatacterization, extend this class.

Let $\mathcal{U}$ be a class of subsets of Θ satisfying A_1, A_2 and A_3 of section 3 and let α be a function from $\mathcal{U}$ to $[0,1]^\Theta$ such that $\Sigma\{\alpha(U) : U \in \mathcal{U}\} = e$ and $\alpha_\theta(U) = 0$ when $\theta \notin U \in \mathcal{U}$. We define the experiment $\mathcal{U}_\alpha$ in the obvious way i.e.; $\mathcal{U}_\alpha = ((\mathcal{U}, \text{all subsets}), (Q_\theta : \theta \in \Theta))$ where Q_θ assigns, for each $\theta \in \Theta$ and each $U \in \mathcal{U}$, mass $\alpha_\theta(U)$ to U. These experiments generalizes the experiments $\mathcal{J}_a$ since $\mathcal{J}_a = \mathcal{U}_\alpha$ where $\mathcal{U} = \{U : U \subsetneqq \Theta \text{ and } \# U = 1 \text{ or } = \#\Theta\}$ and $\alpha_\theta(\Theta) = a_\theta$; $\theta \in \Theta$.

By A_2 the restriction of $\mathcal{U}_\alpha$ to $U \in \mathcal{U}$ is equivalent to $\mathcal{J}_{\{\alpha_\theta(U) : \theta \in U\}}$. The product $\mathcal{U}_\alpha \times \mathcal{U}'_{\alpha'}$ is equivalent to $\mathcal{U}''_{\alpha''}$ where $\mathcal{U}'' = \{U \cap U' ; U \in \mathcal{U}, U' \in \mathcal{U}'\}$ and

$$\alpha''_\theta(U'') = \Sigma\{\alpha_\theta(U)\alpha_\theta(U') : U \in \mathcal{U}, U' \in \mathcal{U}', U \cap U' = U''\} \quad \text{when} \quad U'' \in \mathcal{U}''.$$

It is straight forward to check that $\mathcal{U}''$ satisfies A_1 and A_2. The proof of the fact that $\mathcal{U}''$ also satisfies A_3 is a bit more involved. This, however, follows directly from corollary 3.5 and our next and conclusive theorem.

<u>Theorem F.5.5</u>

An experiment $\mathcal{E}$ permits the implication "$\mathcal{E} \leqq \mathcal{F} \Rightarrow \mathcal{E} \mid \mathcal{F}$" if and only if $\mathcal{E}$ is equivalent to one of the experiments $\mathcal{U}_\alpha$.

Consider the experiments $\mathcal{F} = ((\mathcal{Y}, \mathcal{S}), (Q_\theta : \theta \in \Theta))$ and $\mathcal{U}_\alpha$.

The following conditions are equivalent:

(i) $\mathcal{U}_\alpha \leqq \mathcal{F}$

(ii) $\mathcal{U}_\alpha \mid \mathcal{F}$

(iii) $\alpha_\theta(U) \geqq \|$the $\Sigma\{Q_{\theta'} : \theta' \in U, \theta \neq \theta\}$ absolutely continuous part of $Q_\theta\|$ when $\theta \in U \in \mathcal{U}$, and $\underset{F}{\wedge} Q_\theta = 0$ when F is not contained in any $U \in \mathcal{U}$.

<u>Proof</u>: The equivalence of the three conditions will, since (ii) $\Longrightarrow$ (i) is trivial, follow if we can prove (i) $\Longrightarrow$ (iii) and (iii) $\Longrightarrow$ (ii).

(i) $\Longrightarrow$ (iii): Follows from the remark after corollary 4.4. and, since $[\mathcal{U}_\alpha]_U \sim \mathcal{S}_{\{\alpha_\theta(U) : \theta \in U\}}$, from corollary 5.4.

(iii) $\Longrightarrow$ (ii): This follows from the corollaries 4.4. and 5.4.

It remains to prove the "only if". Suppose the experiment $\mathcal{E}$ with standard measure S permits the implication "$\mathcal{E} \leqq \mathcal{F} \Longrightarrow \mathcal{E} \mid \mathcal{F}$". Put $\mathcal{U} = \{\Theta_\xi : \xi \in \operatorname{supp} S\}$. Then, since $\int x_\theta \, S(dx) \underset{\theta}{\equiv} 1$, $\mathcal{U}$ satisfies A_1. It follows from proposition 3.7 and corollary 3.9 that $\mathcal{U}$ satisfies, respectively, A_2 and A_3. Hence $\mathcal{E} \sim \mathcal{U}_\alpha$ where $\alpha_\theta(\Theta_\xi) = S(\xi)\xi_\theta$; $\theta \in \Theta$, $\xi \in \operatorname{supp} S$. $\square$

F.6 References

[1] Blackwell, D. (1951), Comparison of experiments. Proc. Second Berkeley Symp. Math. Statist. Prob. University of California Press.

[2] Heyer, H. (1969), Erschöpftheit und Invarianz beim Vergleich von Experimenten. Z. Wahrscheinlichkeitstheorie verw. Geb. 12, 21-55.

[3] Le Cam, L. (1964), Sufficiency and approximate sufficiency. Ann. Math. Statistics 35, 1419-1455.

[4] Strassen, V. (1965), The existence of probability measures with given marginals. Ann. Math. Statistics 36, 423-439.

[5] Torgersen, E.N. (1970), Comparison of experiments when the parameter space is finite. Z. Wahrscheinlichkeitstheorie verw. Geb. 16, 219-249.